Attorney Docket No. CISCP718C1

APPLICATION FOR UNITED STATES PATENT

BLOCK CODING SYSTEM AND METHOD FOR DATA AGGREGATION

By Inventors:

Tuchih Tsai
3526 Oxford Drive
Santa Clara, CA 95051
Citizen of the United States

Rishy Mathew
2250 Monroe Street
Apt. #366
Santa Clara, CA 95050
Citizen of India

Assignee: Cisco Technology, Inc.
170 W. Tasman Drive
San Jose, CA 95134

Entity: Large

RITTER, LANG & KAPLAN LLP
12930 Saratoga Ave., Suite D1
Saratoga, CA 95070
(408) 446-8690

BLOCK CODING SYSTEM AND METHOD FOR DATA AGGREGATION

BACKGROUND OF THE INVENTION

The present invention relates generally to communication networks, and more specifically, to a coding scheme for data aggregation of 8b/10b protocols.

Wavelength Division Multiplexing (WDM) is an optical technology that couples many wavelengths in the same fiber, thus effectively increasing the aggregate bandwidth per fiber to the sum of the bit rates of each wavelength. Dense WDM (DWDM) is a technology with a larger (denser) number of wavelengths (e.g., >40) coupled into a fiber than WDM. Systems may support, for example, 100 wavelengths per fiber, enabling a single fiber to carry several hundred gigabits of information. DWDM increases the capacity of embedded fiber by first assigning incoming optical signals to specific frequencies within a designated frequency band and then multiplexing the resulting signals out onto one fiber. DWDM combines multiple optical signals so that they can be amplified as a group and transported over a single fiber to increase capacity. Each signal can be at a different rate and in a different format. DWDM

applications include ultra-high bandwidth long haul as well as ultra-high-speed metropolitan or inner city-networks, and at the edge of other networks such as SONET, Internet protocol (IP) and asynchronous transfer mode (ATM).

In conventional DWDM applications, many different protocols need to be aggregated into a single data channel to be transported over a wavelength. For example, a single 10 Gigabit DWDM channel can carry multiple lower-speed source application protocols such as ESCON (200M), Fiber Channel (1.0625G) or Gigabit Ethernet (1.35G), and multiple channels of those protocols need to be converged to one transport interconnect signal, such as the 10 Gigabit Ethernet (IEEE 802.3.ae) or SONET (OC-48, OC-192), which are the most prominent choices for DWDM technology.

Communication systems transmit signals by means of a number of coding techniques. The 8b/10b code translates 8 bits into one of 256 predetermined 10-bit codes. The extra two bits of code serve a variety of purposes, including differentiating control blocks from data blocks, error detection (disparity control), frame delimiting with data transparency, DC voltage balance, and transitions for clock recovery. An 8b/10b encoder is used to convert 8b data into 10b data. 8b/10b encoding is used in transmitting data on Fibre Channel, ESCON (Enterprise Serial Connection), and Gigabit Ethernet, for example. ESCON was developed by IBM and is widely deployed in large mainframe installations. A

mainframe computer may have hundreds of ESCON I/O channels. These channels may be connected to other mainframes or to peripheral devices. Fibre channel is a standard used for the same set of applications as ESCON.

Fig. 1 illustrates operation of the 8b/10 code. The code combines a 5b/6b code and a 3b/4b code. A mapping is defined that maps each of the possible 8-bit source blocks into a 10-bit code block. There is also a function called disparity control that keeps track of the excess of zeros over ones or ones over zeros. The unencoded input includes 8 bits (A, B, C, D, E, F, G, H) and a control variable. The control line input indicates whether lines A through H are data or control bits. The input information is encoded into bits (a, b, c, d, e, f, g, h, i, j) of a 10-bit transmission character. The control variable is either a "D" for data characters or a "K" for special characters. Each valid transmission character is given a name using the following convention: Kxx.y, where K is the control variable of the unencoded information byte, xx is the decimal value of the binary number composed of the bits E, D, C, B, and A, and y is the decimal value of the binary number composed of the bits H, G of the unencoded information byte. The information received is recovered 10 bits at a time and those transmission characters used for data are decoded into one of the 256 8-bit combinations. Some of the remaining transmission characters are used for protocol management functions.

Since the DWDM protocol of either 10 Gigabit Ethernet or SONET is different from the source application protocols, one way to converge the application protocols is to participate the protocol and terminate the application at the connecting node and regenerate the application protocol at the other end. This, however, involves complicated and expensive system designs. A more efficient method is to perform a transparent interconnect, so that the application protocol data is directly carried through 10 Gigabit Ethernet or SONET. However, protocols such as ESCON and Fibre Channel use the 10b code of the 8b/10b scheme as a part of the protocol signaling. The 10 Gigabit Ethernet employs 8b/10b code for physical communication; however, it does not carry the 10b symbols on user data. SONET uses a total different scrambling scheme and can only carry 8b data.

There is, therefore, a need for a method and system for translating source 8b/10b protocol signals into 8-bit user data.

SUMMARY OF THE INVENTION

A method of coding data for communication within a network is disclosed. The method generally includes receiving an 8b/10b source protocol data stream containing 10-bit code and translating the data stream into an 8-bit code by converting the 10-bit code into 8-bit data and an ordered set. The method further includes transmitting the 8-bit code.

In another aspect of the invention, a method of aggregating different protocols into a single data channel generally comprises receiving a source protocol data stream comprising a plurality of input bits containing data and control information. The method further includes encoding the input bits to translate the data stream into a code containing bit data having a smaller number of bits than the code of the received signals, wherein the encoded bits contain data and ordered sets. The encoded code is combined with another data stream of encoded code having a different source protocol.

A system for coding data for communication within a network generally comprises an ordered set processor operable to receive an 8b/10b source protocol data stream containing 10-bit code and translating the data stream into an 8-bit code by converting the 10-bit code into 8-bit data and an ordered set. The system further includes a data encoder/decoder.

A computer program product for coding data for communication within a network generally comprises code that receives an 8b/10b source protocol data stream containing 10-bit code; code that translates the data stream into an 8-bit code by converting the 10-bit code into 8-bit data and an ordered set; and code that transmits the 8-bit code. A computer-readable storage medium is provided for storing the codes.

The above is a brief description of some deficiencies in the prior art and advantages of the present invention. Other features, advantages, and embodiments of the invention will be apparent to those skilled in the art from the following description, drawings, and claims.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1 is a diagram illustrating a prior art 8b/10b coding scheme.

Fig. 2 is diagram illustrating link connections of the present invention in an ESCON/Fibre Channel application.

Fig. 3 is a diagram illustrating details of the link connection of Fig. 2.

Fig. 4 is a table showing fields for a header symbol of the present invention.

Fig. 5 is a table showing length field.

Fig. 6 is a table showing symbol block encoding.

Fig. 7 is a table showing management symbol encoding.

Fig. 8 is a table illustrating an example of an ESCON data encode.

Fig. 9 is a table listing ESCON control ordered sets.

Fig. 10 is a table listing Fibre Channel control ordered sets.

Fig. 11 is a system block diagram of a computer system that can be utilized to execute software of an embodiment of the present invention.

Corresponding reference characters indicate corresponding parts throughout the several views of the drawings.

DETAILED DESCRIPTION OF THE INVENTION

The following description is presented to enable one of ordinary skill in the art to make and use the invention. Descriptions of specific embodiments and applications are provided only as examples and various modifications will be readily apparent to those skilled in the art. The general principles described herein may be applied to other embodiments and applications without departing from the scope of the invention. Thus, the present invention is not to be limited to the embodiments shown, but is to be accorded the widest scope consistent with the principles and features described herein. For purpose of clarity, details relating to technical material that is known in the technical fields related to the invention have not been described in detail.

The present invention operates in the context of a data communication network including multiple network elements. A system of the present invention provides an efficient, low-latency and high-density scheme to aggregate 8b/10b based protocols over a DWDM network. The network may be configured to use protocols such as ESCON, Fibre Channel or Gigabit Ethernet, for example. The system translates source 8b/10b protocol signals into 8-bit user data.

Referring now to the drawings, and first to Fig. 2, a link connection 20 of the present invention may provide interconnection for 200 Mb ESCON and 1.0625Gb Fibre Channel to 1.25 Gigabit Ethernet on a user 1 Gigabit medium independent interface (GMII), for example. The connection is transparent to users, applications, and other upper layer protocols of the ESCON and Fibre Channel links. An application or device 22 sends data over an ESCON or Fibre Channel link to a Gigabit Ethernet channel that can carry multiple lower-speed source application protocols. The 10-bit code of the 8b/10b scheme is used by ESCON and Fibre Channel as part of the protocol signaling. Since the 10-bit control characters are a part of the protocols, they must be carried through as well as the user data. Although the Gigabit Ethernet employs the 8b/10b code, it does not use any 10-bit control symbols in the protocol. The link connection translates the source protocol signals into an 8-bit code for use by the Gigabit Ethernet protocol.

Fig. 3 illustrates additional detail of the link connection 20 of Fig. 1. The 10-bit of ESCON and Fibre Channel is received at the link connection 20, processed by an ordered set processor 30,and encoded by a data encoder/decoder 32 to create 8-bit code for the Gigabit Ethernet interface. The ordered set processor 30 uses an 8-bit code to efficiently map both K and D data, as further described below. The data encoder 32 is used to encode 10-bit symbols of ESCON and Fibre Channel into the 256 domain of 8-bit code. The decoder may

perform in a cut-through byte by byte operation since all information is encoded in sequence. When an error is encountered, the decoder can implement a local error reporting and accounting method.

The 8b/10b code provides a potential 1024 code words for signaling. A convention of K/D with numbers is used to indicate the data. The 8-bit 256 user data is represented by Dxx.y, other control codes are represented by Kxx.y where xx is from 0 to 31 and y is from 0 to 7. For example, D28.5 represents 10b code 0011110010 or 1100000010 depending on the code disparity.

The 8b/10b based protocol uses a 10-bit domain data as a part of the protocol signaling called ordered set. An ordered set is a 4-byte transmission word which has a special character as its first transmission character. An ordered set may be a frame delimiter, a primitive signal, or a primitive sequence. Ordered sets are used to distinguish control information from data. For example, in ESCON, K28.5 itself is used for IDLE, K28.1 + K28.7 is Start of Frame (SOF), K28.6 + K28.2 + K28.5 is End of Frame (EOF), and ABORT is K28.6 + K28.4 + K28.4. In Fibre Channel, the ordered set has four symbols. For example, IDLE consists of K28.5 + D21.4 + D21.5 + D21.5. The ordered set may be any number and combination of K and D symbols.

The present invention uses an 8-bit code to efficiently map both K and D data. The system uses a protocol symbol that is a K control character. The K

control character may be followed by one to three user data codes (e.g., Dxx.y), or not followed by any D codes. An 8-bit domain data block referred to as a frame, encapsulates a source of 8b/10b protocol data stream (i.e., stream of K and D data of the source protocol) and is sent through a DWDM channel. The length of the frame is determined by bandwidth factors between the total number of the source protocol channels and the DWDM data rate and is preferably implementation and system dependent.

A frame starts with a one byte header to begin the encapsulation and also provides a 5-bit sequence number for the frame. The header identifies the beginning of a mapped data frame and is the first byte of the encapsulation packet. The header also identifies whether the character following the header is a symbol (K) or data (D). Details of the header symbol are shown in Fig. 4. The header sequence count starts from zero after reset and is then continuously incremented. The sequence number is used for error detection at a receiving node.

As previously discussed, the encoding scheme of the present invention uses an 8-bit code to represent a K character that may be followed by up to 3 bytes of D data characters. More than 3 bytes of D data characters are preceded with a length field that points to the end of a data block. The data length field indicates the number of user data between symbols. Due to latency

considerations, the actual length number may vary. The minimum number of length is one byte, as shown in the table of Fig. 5. If the length is two bytes or more, the length count continues and Bits 5-0 are the MSBs (most significant bytes) of the total length. In each realization, the encoder and decoder use the same number length. An 8b/10b code error in a data block is indicated once in its data length field until a correct K character is observed or a new data length field is established to guarantee the encoding bandwidth.

After a length indicated D data block, either another symbol code or data length field may follow. The data length field proceeds a data block for data more than 3 bytes, indicating the length of the data as well as whether the character that follows the data block is a symbol (K) or another data length field. The procedure continues and can stop at any time to complete a frame. The data integrity of the frame is protected using the DWDM data protocol's protection code (e.g., 32-bit CRC in 10G Ethernet).

The block coding also includes a symbol encode, which is an 8-bit data encoding the K and D combination, and a management encode, which is an 8-bit data encoding management messages. The symbol block encoding is shown in Fig. 6. The protocol symbol K character is encoded in 8-bit data and indicates if the character following the symbol block is another symbol (K) or data (D). The data length field follows data (D). The symbol encode includes all of the assigned

ordered sets. Bit 7 indicates if the beginning of the next block is a symbol. Bits 6 and 5 indicate how many data bytes follow the symbol as a part of the ordered set. Bits 4 to 0 are symbol code and identify the control character of the symbol. Bit 4=0 is used for well-known control code, and bit 4=1 is used for other management messages such as error. For example, in Fibre Channel, the symbol is x 1110101 which is K28.5 followed by 3 bytes of data as a part of the ordered set.

The management symbol encode is shown in Fig. 7. Management messages such as error are defined in the table of Fig. 7. An error symbol indicates that there was an error when receiving the character which can not be identified. A null symbol is used for data filling and is discarded by the receiving decoder. The current encapsulated packet is terminated by EOF (end of field) symbol and the rest of the data is discarded by the receiving decoder. An 8b/10b code error from the source protocol is encoded in a management encode for error unless the error starts in a data block when a data length is already established. After a reset occurs, the decoder assumes the expecting sequence count to be zero and replaces its value with the received packet sequence count after comparing them. An error shall be reported if the expecting sequence count does not match with the packet sequence count.

Fig. 8 illustrates an example of an ESCON data encode. The system described above allows the source protocol data stream to be segmented at any length and any time, providing full flexibility for other design considerations such as bandwidth factor. For example, the frame length is different in encapsulating 40 ports of ESCON and 10 ports of Fibre Channel, however, the same scheme applies. Also, the coding enables the design to achieve high performance of low latency and high density data aggregation. The scheme of the present invention may be implemented in a staged cut-through pipeline operation to further reduce encoding latency since it does not need to complete the entire frame before it starts to send it. The encoding scheme may be applied to the 8b/10b based protocols and the decoder can operate entirely in cut-through mode.

Figs. 9 and 10 show ordered sets used in the 8b/10b protocol signaling of ESCON and Fibre Channel, respectively. The tables include a list of functions and the corresponding ordered sets. The running disparity is also included in the table containing the Fibre Channel ordered sets. In ESCON the number of K or D code in an ordered set is not fixed, while in Fibre Channel, the ordered set starts with K28.5 and is followed by 3 data (D) codes.

It is to be understood that the block coding scheme of the present invention may be used with protocols other than ESCON and Fibre Channel, without departing from the scope of the invention. Also, the data fields and

ordered sets shown and described above are provided only as examples and other formats and ordered sets may be used without departing from the scope of the invention. Furthermore, the system and method of the present invention may be used for data aggregation of coding protocols other than 8b/10b.

Fig. 11 shows a system block diagram of computer system 38 that may be used to execute software of an embodiment of the invention. The computer system 38 includes memory 42 which can be utilized to store and retrieve software programs incorporating computer code that implements aspects of the invention, data for use with the invention, and the like. Exemplary computer readable storage media include CD-ROM, floppy disk, tape, flash memory, system memory, and hard drive. Additionally, a data signal embodied in a carrier wave (e.g., in a network including the Internet) may be the computer readable storage medium. Computer system 38 further includes subsystems such as a central processor 40, fixed storage 44 (e.g., hard drive), removable storage 46 (e.g., CD-ROM drive), and one or more network interfaces 54. Other computer systems suitable for use with the invention may include additional or fewer subsystems. For example, computer system 38 may include more than one processor 40 (i.e., a multi-processor system) or a cache memory. The computer system 38 may also include a display, keyboard, and mouse (not shown) for use as a host.

The system bus architecture of computer system 38 is represented by arrows 56 in Fig. 11. However, these arrows are only illustrative of one possible interconnection scheme serving to link the subsystems. For example, a local bus may be utilized to connect the central processor 40 to the system memory 42. Computer system 38 shown in Fig. 11 is only one example of a computer system suitable for use with the invention. Other computer architectures having different configurations of subsystems may also be utilized. Communication between computers within the network is made possible with the use of communication

As can be observed from the foregoing, the present invention has numerous advantages. The present invention provides a system that can achieve high performance of low latency and high density of data aggregation. The scheme may be implemented in a staged cut-through pipeline operation to further reduce encoding latency since it does not need to complete the entire frame before starting to send it.

Although the present invention has been described in accordance with the embodiments shown, one of ordinary skill in the art will readily recognize that there could be variations made to the embodiments without departing from the scope of the present invention. Accordingly, it is intended that all matter contained in the above description and shown in the accompanying drawings shall be interpreted as illustrative and not in a limiting sense.